

Application No. 09/683,235

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

B1
Claim 1 (Currently Amended): A method for automatically generating a query from selected document content, comprising:

defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;

automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;

automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and

automatically formulating a the query that restricts to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.

Claim 2 (Original): The method according to claim 1, further comprising limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.

Claim 3 (Original): The method according to claim 2, wherein the number of terms added is limited to a predefined number.

Application No. 09/683,235

B1
Claim 4 (Original): The method according to claim 2, further comprising limiting the query by adding terms defining the assigned classification label.

Claim 5 (Original): The method according to claim 1, wherein the organized classification of document content is defined using a hierarchical organization.

Claim 6 (Original): The method according to claim 1, further comprising using a text categorizer to assign the classification label assigned from the organized classification of content.

Claim 7 (Original): The method according to claim 6, further comprising:
extracting with the text categorizer a set of terms relating to the document content; and
appending to the query ones of the set of terms extracted by the text categorizer to contextualize the query.

Claim 8 (Currently Amended): The method according to claim 7, further comprising abbreviating the set of terms extracted by the text categorizer to a predefined ~~limit of terms~~ number of terms.

Claim 9 (Original): The method according to claim 8, wherein said abbreviating comprises:

extracting noun phrases from the selected document content;
ranking the noun phrases by those that occur most frequently in the document content;
defining a subset of noun phrases by identifying those ranked noun phrases that occur more frequently than a first predefined frequency;
ranking those words in the subset of noun phrases by their frequency of occurrence to define an ordered list of words;

Application No. 09/683,235

B/ defining a subset of the ordered list of words by identifying those ranked words that occur more frequently than a second predefined frequency;

re-ranking the subset of words in inverse frequency to their use in the category of information in the information retrieval system identified by the assigned classification label;

using only those highest ranked words in the re-ranked subset of words to define the set of terms appended to the query.

Claim 10 (Original): The method according to claim 1, wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary.

Claim 11 (Original): The method according to claim 10, further comprising ranking results from the query performed at the information retrieval system in accordance with one of the assigned classification label and the characteristic vocabulary.

Claim 12 (Original): The method according to claim 11, using the method in a system for enriching selected content of a document with personalities that identify enrichment themes.

Claim 13 (Original): The method according to claim 1, further comprising automatically identifying the set of entities using a service that recognizes entities of a predefined type.

Claim 14 (Currently Amended): A system for automatically generating a query from selected document content, comprising:

an entity extractor for automatically identifying a set of entities in the selected document content for searching information related thereto using an information retrieval system;

Application No. 09/683,235

B1
a categorizer for defining an organized classification of document content with each class in the organization of content having associated therewith a classification label; each classification label corresponding to a category of information in the information retrieval system; the categorizer automatically assigning the selected document content a classification label from the organized classification of content; and

a query generator for automatically formulating a the query that restricts to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.

Claim 15 (Original): The system according to claim 14, further comprising a short length aspect vector generator for generating terms relating to context information surrounding the set of entities in the selected document content; wherein the query generator adds the terms relating to the context information to limit the query.

Claim 16 (Original): The system according to claim 15, wherein the query generator further limits the query by adding terms defining the selected classification label provided by the categorizer.

Claim 17 (Original): The system according to claim 16, further comprising a content manager for enriching the selected document content with results provided from the information retrieval system using the query.

Claim 18 (Currently Amended): An article of manufacture for use in a computer system, comprising:

a memory;

instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:

Application No. 09/683,235

B1
defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;

automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;

automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and

automatically formulating ~~a the query that restricts to restrict~~ a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label.

Claim 19 (Original): The article of manufacture according to claim 18, wherein the instructions stored in the memory further comprise limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content.

Claim 20 (Original): The article of manufacture according to claim 19, wherein the instructions stored in the memory further comprise further limiting the query by adding terms defining the assigned classification label.